**Supplemental Text**

 *Measuring Dynamic State Public Opinion via MRP*

One technique which has been shown to overcome the challenges in measuring state opinion over time from national surveys is multilevel modeling, imputation, and post-stratification (referred to as MRP) developed by Gelman and Little (1997) and extended by Park et al. (2004; 2006), Lax and Phillips (2009), and Pacheco (2011). First, we begin with a multilevel model to estimate state public opinion for individuals given demographics and state. The MRP approach includes various predictors to estimate state public opinion. Following Pacheco (2011), I use gender (0=male, 1=female), race (0=non-black, 1=black), age (four categories: 18-29, 30-44, 45-64, and 65+) and education (four categories: no high school degree, high school degree, some college, and college+) for estimating public opinion towards both education and welfare. I write the model below using indexes *j,k,* and *l* for state, age category, and education category, respectively; the subscript *i* refers to individual respondents.[[1]](#footnote-1)

(1) **Level 1:** Pr(*yi*=1) = logit-¹( β0 + β1Femalei + β2Blacki + αj[i] + αk[i] + αl[i])

(2) **Level 2:** αj ~ N (0, σ²state) for *j*=1,…,51

 αk ~ N (0,σ²age) for *k*=1,…,4

 αl ~ N (0,σ²education) for *l*=1,…,4

The next step is imputation. Like Pacheco, I define each combination of demographic characteristics and state (for instance, a non-black, female, aged 18-29, with a high school degree from Connecticut) as a “person type.” Each of the 3,264 person types has an associated probability of supporting a particular policy, which is modeled in the multilevel regression as a function of state, gender, age, race and education. Imputation is conducted on each person type even if absent from the sample. After imputation, we have *θc*, which is the inverse logistic given the relevant predictors and their estimated coefficients (*θc*, is an average based on 1,000 simulations with *c* indexing the 3,264 unique combinations).

The final stage is post-stratification. Post-stratification corrects for differences between state samples and state populations by weighting the predicted values of each person type in each state by actual Census counts of that person type in a state. For example, the 2000 Census reports that there were 581 people who were white, male, age 18-29, no high school degree, and living in Alabama: 1.7% of Alabama’s population. The imputed opinion of each person type, *θc,* is then weighted by the corresponding population frequencies. In the final step, we calculate the average response over each person type in each state and summarize to get point predictions and uncertainty intervals.

*Adding a Time Component*

As suggested by Pacheco, I add a time component by pooling surveys across a small time frame. For both education and welfare, I employ five year moving averages, pooling individual responds on surveys from the specified time. For instance, to get point estimates for 1979 using a five year pooled window, I combine estimates from 1977, 1978, 1979, 1980, and 1981 and then perform the MRP technique on this pooled dataset. It is important to note that the five year window may not include consecutive years if the public opinion measure was not asked in a particular year. For instance, since all states are missing on both public opinion measures for 1995, 1994 estimates are obtained from pooling surveys from 1992, 1993, 1994, 1996, and 1998.

The MRP process is repeated for each year after moving the time frame up a year at a time. By pooling and taking the median year, the first two and last two years are missing for the five year window. Hence, for attitudes towards education, I have yearly state estimates from 1975-2000 and for attitudes towards welfare, I have yearly state estimates from 1974-2000. Pacheco shows that while there is a tradeoff between the reliability of estimates and sensitivity to very short-term shocks, the efficiency benefits of pooling over a small time period outweighs the costs of biasedness.

*House Effects or Question Wording*

 I depart from Pacheco (2011) in that I combine survey responses across various polling organizations to increase the time series data and reliability of the measures (see Table A1 for specific question wording). This opens up the possibility that the observed dynamics are artificial, resulting from “house effects” or question wording. The five year average deflates the possibility that house effects are causing changes in one particular year. I also only use survey questions asked in identical ways to minimize the effects of question wording. Finally, as shown in Tables A2 and A3, a comparison of frequencies across survey organizations in particular years shows that there are minimal differences across organization. Thus, I am confident that changes in education or welfare preferences across the states are not artificially caused by survey differences.



Table A2 Raw Frequencies of Education Spending Opinion across Survey Organizations



Table A3 Raw Frequencies of Welfare Spending Opinion across Survey Organizations



Table A4. Descriptive Statistics for All Variables



1. Following Park et al. (2004; 2006) and Gelman and Hill (2007), I fit the model using the Bayesian software *WinBugs* (Spiegelhalter et al. 1999) as called from R (R Development Core Team 2003) using Gelman’s (2003) *Bugs.R.* Bayesian multilevel models are especially useful for more complicated multilevel models, for example those with non-nested components, and also allow the estimation of uncertainty by using prior distributions, which are given to all parameters (Gelman & Hill 2007 345). Parameters can then be drawn from these distributions over a number of simulations. I assign normal distributions to the coefficients with means of 0 and standard deviations σ²state, σ²age, σ²educ, estimated from the data given non-informative uniform prior densities (Park et al. 2004 378). [↑](#footnote-ref-1)